



Research Department Report

The visibility of noise in System I colour television

A. Oliphant M.A., C.Eng., M.I.E.E., K.J. Taylor, N.T. Misson

THE VISIBILITY OF NOISE IN SYSTEM I PAL COLOUR TELEVISION

A. Oliphant, M.A., C.Eng., M.I.E.E., K. J. Taylor, N. T. Misson

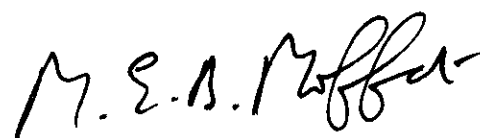
Summary

The results of new subjective tests on the visibility of noise in System I PAL colour television are reported. Three forms of noise are considered:

*white (flat) noise,
de-emphasised triangular noise,
and the noise resulting from vestigial sideband demodulation (VSB noise).*

The results are compared with those given by work published 20 years ago and it is concluded that observers today are more critical of noise in pictures: about 2.5 dB less noise is required to produce the same impairment.

Issued under the Authority of



Research Department, Engineering Division,
BRITISH BROADCASTING CORPORATION

Head of Research Department

THE VISIBILITY OF NOISE IN SYSTEM I PAL COLOUR TELEVISION

A. Oliphant, M.A., C.Eng., M.I.E.E., K. J. Taylor, N. T. Misson

1.	Introduction	1
2.	Review of Previous Work	1
3.	Experimental Details	2
3.1	General	2
3.2	Noise sources	2
3.3	Picture sources	2
3.4	Viewing conditions	4
3.5	Test procedure	4
4.	Results	4
4.1	Results of first series of tests (white noise and de-emphasised triangular noise)	4
4.1.1	Effect of viewing distances	4
4.1.2	Effect of expert and non-expert viewers	5
4.1.3	Effect of noise type on the pictures used	5
4.1.4	Mean results	5
4.2	Results of the second series of tests	6
4.3	Comparison of the results of the two series of tests	7
4.3.1	Consistency of results for white noise and de-emphasised triangular noise ...	7
4.3.2	Comparison of results for white noise, VSB noise and de-emphasised triangular noise	7
4.4	Further analysis of results	8
5.	Comparison with Other Results	8
5.1	Comparison with the results of Allnatt and Prosser	8
5.2	Comparison with the reference impairment characteristic of CCIR Recommendation 654	10
5.3	Comparison with results derived from weighting curves	10
6.	Recommendations	11
7.	Conclusions	11
8.	References	12
	Appendix 1: The Effect of the PAL Delay Line on Chrominance Signal-to-noise Ratio	13
	Appendix 2: An Experiment on the Relative Visibility of Luminance and Chrominance Noise	14

© BBC 2003. All rights reserved. Except as provided below, no part of this document may be reproduced in any material form (including photocopying or storing it in any medium by electronic means) without the prior written permission of BBC Research & Development except in accordance with the provisions of the (UK) Copyright, Designs and Patents Act 1988.

The BBC grants permission to individuals and organisations to make copies of the entire document (including this copyright notice) for their own internal use. No copies of this document may be published, distributed or made available to third parties whether by paper, electronic or other means without the BBC's prior written permission. Where necessary, third parties should be directed to the relevant page on BBC's website at <http://www.bbc.co.uk/rd/pubs/> for a copy of this document.

THE VISIBILITY OF NOISE IN SYSTEM I PAL COLOUR TELEVISION

A. Oliphant, M.A., C.Eng., M.I.E.E., K. J. Taylor, N. T. Misson

1. INTRODUCTION

Analogue signals are always accompanied by noise. The acceptable level of noise is an important parameter in broadcasting: each component of the broadcast chain adds some noise to the signal and the service area of the final transmitter is limited by receiver noise. The visibility of noise in television has therefore been studied since the early days of the television service; Refs. 1-6 all report experiments using modern subjective test methodologies, and Weaver¹ gives references going back to 1939.

The question of the visibility of noise has recently been raised again by the proposal for digital stereo sound with terrestrial television. In setting the parameters of the digital signal, it is important to ensure that it is not affected by noise at the limit of the service area for the vision signal. A review of previous work was undertaken and some new work was carried out to clarify the relationship between vision carrier-to-noise ratio and picture signal-to-noise ratio⁷. At the same time it was decided to investigate again the relationship between picture signal-to-noise ratio (SNR) and subjective impairment: as explained in Section 2, it was felt that the results of previously-published work may not be applicable in today's conditions.

A series of subjective tests was therefore carried out in 1986 using various levels of white noise and de-emphasised triangular noise added to baseband video signals. It was later pointed out that the form of noise that is of interest in planning the service areas of transmitters arises from vestigial-sideband demodulation of a television signal (VSB noise). This form of noise has a spectrum which is flat above 1.25 MHz but is 3 dB lower at low frequencies⁷. Because high frequency noise is less visible than low frequency noise, it would be expected that VSB noise would cause less impairment than white noise for the same signal-to-noise ratio. Previously published work has not considered VSB noise (its effects have been inferred from results for white noise by applying an assumed weighting factor); a second series of tests was therefore arranged to investigate VSB noise, with a few values of white noise and de-emphasised triangular noise included for comparison with the first series of tests. This comparison showed very close consistency between the two series of tests (see Section 4.3), allowing the results of the two series of tests to be compared directly. This Report gives the results of both series of tests and compares them with previously published work.

2. REVIEW OF PREVIOUS WORK

The papers by Weaver¹ and Geddes³ both compare the visibility of noise in 405-line and 625-line monochrome pictures. Their results for 625-line pictures show good agreement, although their results for 405-line pictures differ widely. Barstow and Christopher² carried out tests on the visibility of noise in both monochrome and colour pictures; they concluded that the visibility was the same to within ½ dB of SNR. However, Allnatt and Prosser⁵ compared their colour results with the results of their earlier study of monochrome pictures⁴ and found that colour pictures were less sensitive to noise. In view of this difference it seems prudent to confine attention to the results of tests using colour pictures.

The results most commonly quoted are those of Allnatt and Prosser⁵. This study was published in 1966, before the colour service started in the UK, and used a 625-line version of NTSC rather than the PAL system that was eventually adopted. A later study by White and Reid⁶ gave similar results to those of Allnatt and Prosser, although there were small but significant differences, notably in the visibility of chrominance noise. There were some differences between the two experiments: White and Reid used a higher screen brightness, they used PAL signals rather than 625-line NTSC, and although they used two of the same slides as the earlier experiment, they increased the colour saturation of the signals because they considered the slides deficient in this respect. This increased saturation may account for the increased sensitivity to chrominance noise that they found — chrominance noise is known to be more visible in areas of highly-saturated colour. One would otherwise expect a reduction in the sensitivity to chrominance noise because of the change from NTSC to PAL: the averaging across a line delay in the PAL decoder produces a reduction of 3 dB in the measured chrominance noise in a PAL decoder, although the different appearance of the noise means that the subjective effect is of a reduction of only about 1.0 to 1.5 dB (see Appendix 1).

There are two reasons for expecting that viewers might find a given level of noise more annoying now than twenty years ago. One is that shadowmask display tubes have improved significantly in resolution so that more high-frequency noise actually appears on the display; the other is that viewers' expectations may be higher now that high-quality colour pictures are available daily to almost the whole population, whereas in 1966 colour

television was an exotic novelty. Barstow and Christopher² found such an increase in sensitivity to noise between their experiment in 1962 and a similar experiment in 1953. For these reasons it was thought worthwhile to carry out new tests to obtain an up-to-date assessment of the relationship between video signal-to-noise ratio and subjective impairment.

3. EXPERIMENTAL DETAILS

3.1 General

In general the two series of tests used the same experimental conditions and test methodology. The differences between the two series are summarised in Table 1.

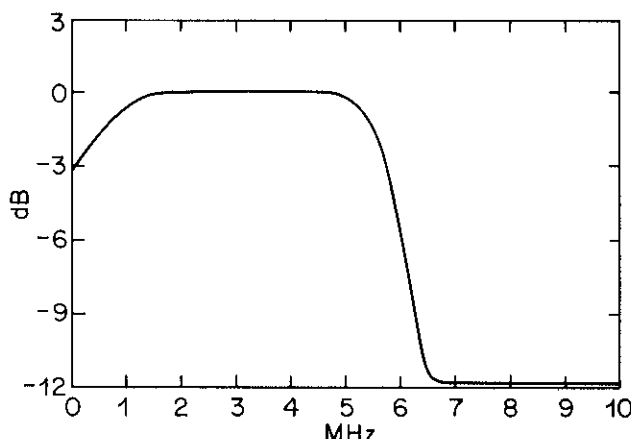


Fig. 1 - Measured frequency spectrum of VSB noise.

Table 1: Summary of differences between test series 1 and 2

	Test series 1	Test series 2
Number and status of observers	9 expert 7 non-expert	20 non-expert
Pictures used	EBU slide 'Girl with Toys' Test Card G (Philips)	EBU slide 'Girl with Toys' Test Card G (Philips) Moving sequence
Conditions and SNRs(dB)		
White noise	24.5, 29, 33.5, 38, 42.5,	28, 38
De-emphasised triangular noise	24, 28.5, 33, 37.5	26.2, 36.2
VSB noise	—	22.2, 27.2, 32.2, 37.2, 42.2
Viewing distances	6H 4H	6H

3.2 Noise sources

White noise was generated by a transistor connected as a reverse-biased diode. The spectrum of the noise was flat to within 1 dB over the range 0 - 5 MHz. To obtain de-emphasised triangular noise, this white noise was amplified and passed through a triangular network followed by a de-emphasis network according to CCIR Recommendation 405.

The VSB noise was produced by an avalanche diode noise source which had a frequency spectrum flat over the range 70 MHz $\pm$ 20 MHz. The output of this noise source was mixed down to television IF frequencies and shaped using a System I VSB receiver filter. After mixing down to baseband frequencies the VSB noise was amplified ready for use in the subjective tests. The measured spectrum of the VSB noise is shown in Fig. 1.

The power of the selected noise source was measured as shown in Fig. 2 by a power meter preceded by a 5.0 MHz lowpass filter as recommended

by the CCIR, while noise outside the 5.5 MHz bandwidth of System I was removed by a 5.5 MHz lowpass filter in the video feed to the monitor.

3.3 Picture sources

Two picture sources were used in the first series of tests: the EBU slide 'Girl with Toys' and Test Card G. The use of test patterns in subjective tests is normally deprecated. However because the presence of noise is not a picture related effect (although its visibility is picture related) the use of a test card was thought to be acceptable. It is a useful picture because it contains areas of plain grey and also areas of fairly saturated colour. 'Girl with Toys', on the other hand, contains a lot of low-saturation colour and has a patterned background.

In the second series of tests the same two still pictures were used along with a moving sequence recorded off air from a live magazine programme on C-format video tape. The SNR of the signal from the videotape recorder was approximately 45 dB. It was

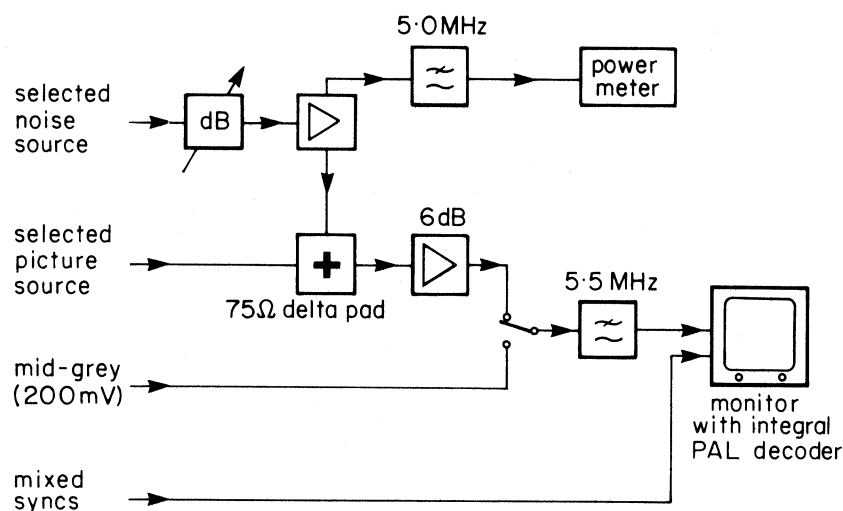


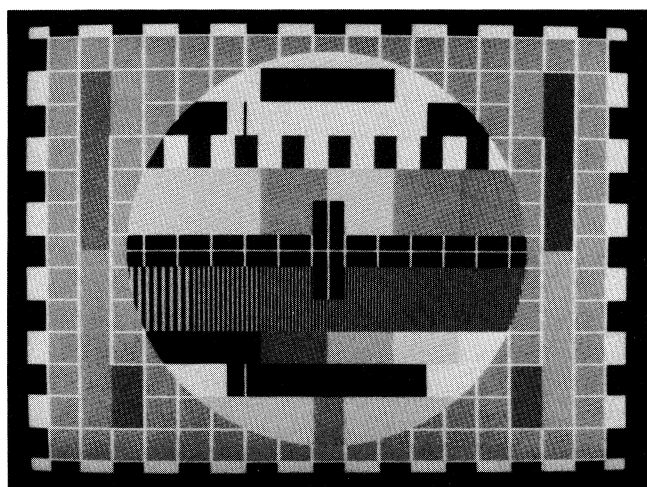
Fig. 2 - Simplified block diagram of the test equipment.

thought that movement in the test material might make noise less annoying (this proved not to be the case — see Section 4.2).

Test Card G was generated by a Philips Test Pattern Generator and 'Girl with Toys' came from a high quality flying-spot slide scanner with electronic masking to correct for the errors in colour rendering which reduce saturation. For convenience in running the tests, both pictures were stored in a digital stills store using the sampling standards of CCIR Recommendation 601. The output of the stills store was PAL coded before adding noise. Fig. 3 shows monochrome versions of the two still pictures used in the tests and a frame from the moving sequence.



(b)



(a)



(c)

Fig. 3 - Monochrome versions of the pictures used in the tests:

- a) Test Card G
- b) Girl with Toys
- c) Still from moving picture sequence

3.4 Viewing conditions

The tests were carried out in a viewing room using a colour monitor with an integral delay-line PAL decoder. The monitor tube had a 500 mm screen diagonal with an in-line gun and delta phosphor dots with a spacing of 0.42 mm. It was positioned in front of a grey wall on a stand draped with grey cloth. Its peak brightness was set to 70 Cd/m² as recommended by the CCIR⁸ and its cutoff was set using the standard lineup waveform (PLUGE). The ambient light level in the test room was low.

In the first series of tests observers were seated at four times and six times picture height (4H and 6H) from the monitor. However, no significant difference was found between the two viewing distances and in the second series of tests all the viewers were seated at 6H.

3.5 Test procedure

The EBU method⁹ was used for the tests. In this method, which is an elaboration of the general principles set out in CCIR Recommendation 500⁸, a reference picture and a test picture are shown in each presentation. At the end of each presentation, the observers are asked to grade the test picture with respect to the reference picture using the CCIR 5-point impairment scale:

Grade	Impairment
5	Imperceptible
4	Perceptible, but not annoying
3	Slightly annoying
2	Annoying
1	Very annoying.

The timing of each presentation was as follows:

Reference test picture	: 10 s
Grey	: 5 s
Impaired test picture	: 15 s
Grey (voting period)	: approx. 20 s

The tests were run and the results analysed using automated equipment. This equipment has been constructed to facilitate the running of subjective tests and the processing of the results*. It consists of hand-held consoles, on which are five buttons marked with the ratings corresponding to various levels of picture impairment. An interface unit connects these consoles to the user port of a BBC Microcomputer, so that software can read in the votes cast by the observers

* The equipment was designed and built by W. J. Collins.

for each picture condition, and store them on a floppy disk. Other software accepts descriptions of pictures and test conditions and creates a pseudo-random display sequence, runs the sequence with prompts for the operator at the correct timings and processes the results obtained, with the correct rejection procedure, all in accordance with EBU Recommendation R28⁹. The results of a test can thus be displayed a few seconds after completing the test.

For both series of tests the observers were divided into groups of four or five. In the first series of tests each group of observers completed the test in one session; in the second series the test was divided into two sessions. In each series, each combination of picture and SNR was shown twice and in each session five extra presentations were added at the start of the session to allow for observer adaptation — the observers graded these extra tests, but the results were ignored by the computer.

4. RESULTS

4.1 Results of first series of tests (white noise and de-emphasised triangular noise)

The use of a computer for registering and analysing the test results allowed a number of aspects of subjective testing to be examined. Comparisons were made on the effect of viewing distance, expert and non-expert observers, noise type, and the two pictures used.

4.1.1 Effect of viewing distances

Only expert observers sat at 4H; both expert and non-expert observers sat at 6H. To eliminate any difference between expert and non-expert observers, results of expert observers at 6H were compared with results of the expert observers at 4H. The results for white noise are shown in Fig. 4. It can be seen that there is no significant difference between 4H and 6H viewers for 'Girl with Toys' and very little difference for Test Card G. (Results for de-emphasised triangular noise — not illustrated — showed more scatter but no significant differences.) These results are difficult to explain. Conventional theory, assuming the angular resolution of the eye does not change with accommodation to different viewing distances, would argue that reducing the viewing distance would increase the effective noise power; going from 6H to 4H should make just under 2 dB difference for flat noise. Although the samples are small (five observers at 6H and four observers at 4H), such a difference ought to be visible. Further work would be needed to explain these results; a more detailed investigation might be

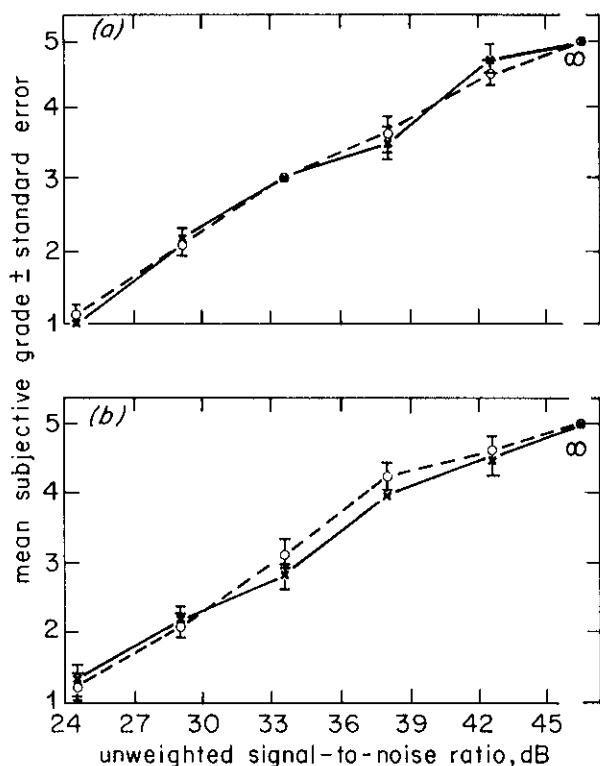


Fig. 4 - Effect of viewing distance:

- a) Girl with Toys
b) Test Card G
○---○ 4 H expert viewers
x—x 6 H expert viewers

worthwhile in the context of HDTV where extrapolations of picture size or viewing distance are often made because of the difficulty of carrying out experiments with full-size HDTV sources and displays.

4.1.2 Effect of expert and non-expert viewers

Fig. 5 shows results separately for expert and non-expert observers, for white noise. As might be expected, both sets of observers agree at the extremes of the scale — unimpaired pictures and badly-impaired pictures — but expert viewers consistently rate about half a grade lower in the middle of the scale. A similar effect was found for de-emphasised triangular noise. In view of these results, it was decided to use only the results for non-expert observers in further comparisons. This is in line with EBU recommendations⁹, and ensures that the results are representative of the viewing population as a whole. Unfortunately, this reduced the sample size from sixteen to seven but, as will be shown in Section 4.3.1, the second series of tests gave a striking

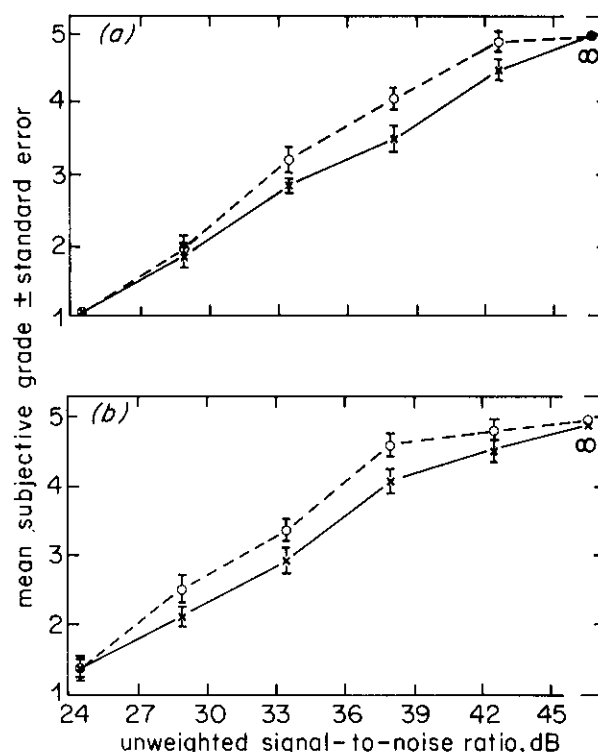


Fig. 5 - Comparison of expert and non-expert viewers' results:

- a) Girl with Toys
b) Test Card G
○---○ Non-expert viewers
x—x Expert viewers

confirmation of the validity of the results.

4.1.3 Effect of noise type on the pictures used

Fig. 6 compares the effect of the two types of noise on the two pictures used. Fig. 6(a) shows that, for a given SNR, white noise is more annoying on 'Girl with Toys' than on Test Card G; Fig. 6(b) shows that the reverse is true for de-emphasised triangular noise. This difference probably arises because de-emphasised triangular noise has a greater high frequency content and therefore gives more chrominance noise — which is more visible in the highly-saturated areas of the test card.

4.1.4 Mean results

Table 2 and Fig. 7 show mean results for the seven non-expert observers. These results are analysed further in Section 4.4. It should be remembered that the results for de-emphasised triangular noise are not truly representative of FM transmission below about 29 dB SNR because threshold spikes would be visible in a real FM transmission.

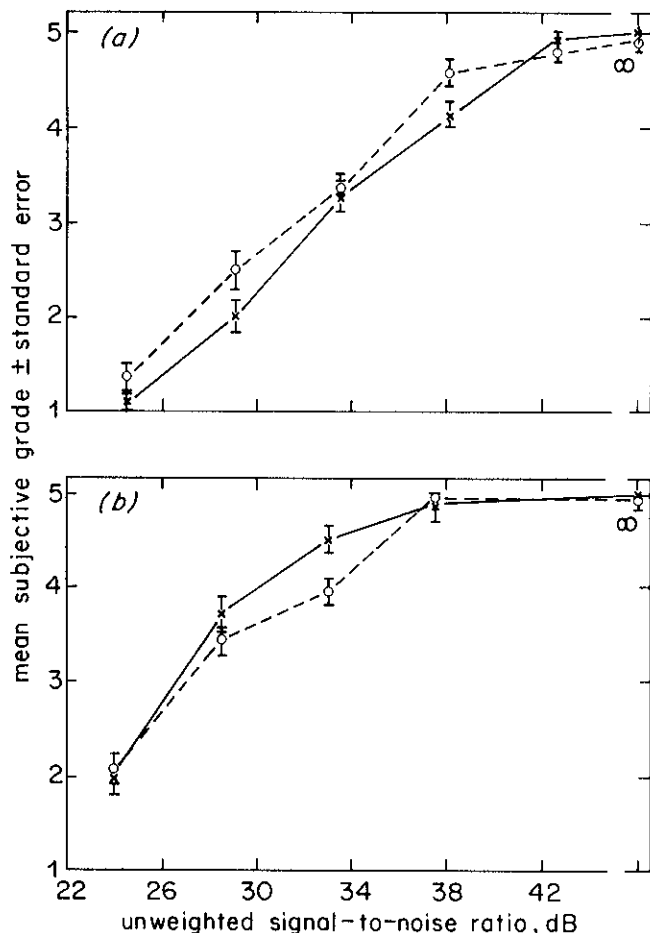


Fig. 6 - Effect of the different noise spectra on the two pictures used:

- a) White noise
b) De-emphasised triangular noise
- — — — ○ Test Card G
× — — — × Girl with Toys

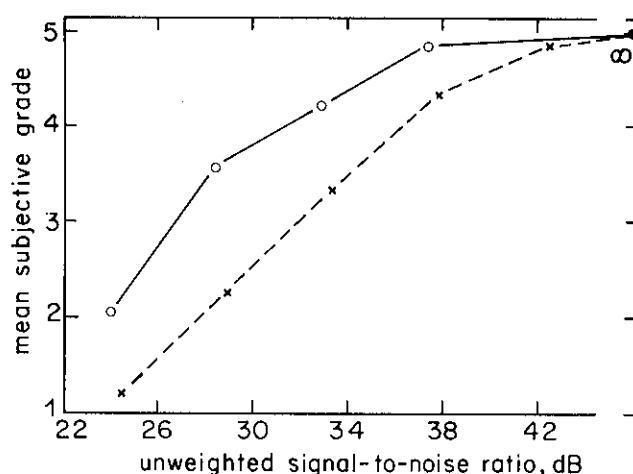


Fig. 7 - Mean results of first series of tests.

4.2 Results of the second series of tests

The second series of tests was intended only to assess the effect of VSB noise; there was no attempt to carry out all the comparisons investigated in the first series. As described in Section 3, twenty non-expert observers took part in the tests, in groups of five. All sat at six times picture height. Five values of VSB noise were used as test conditions, with two values each of white noise and de-emphasised triangular noise for comparison with the first series of tests. Three picture sources were used: 'Girl with Toys' and Test Card G, as before, and also a moving sequence.

Table 2: Results of the first series of tests: results for seven non-expert observers.

Condition		De-emphasised triangular noise				White noise					
Picture	SNR dB	24.00	28.50	33.00	37.50	24.50	29.00	33.50	38.00	42.50	∞
Girl with Toys	$\bar{x}$	2.00	3.71	4.50	4.75	1.07	2.00	3.29	4.14	4.93	5.00
	σ	0.55	0.73	0.52	0.45	0.27	0.68	0.61	0.53	0.27	0.00
	$\sigma/\sqrt{n}$	0.15	0.19	0.14	0.13	0.07	0.18	0.16	0.14	0.07	0.00
Test Card G	$\bar{x}$	2.07	3.43	3.93	4.93	1.36	2.50	3.36	4.57	4.79	4.92
	σ	0.73	0.51	0.47	0.27	0.63	0.71	0.58	0.51	0.43	0.29
	$\sigma/\sqrt{n}$	0.20	0.14	0.13	0.07	0.17	0.22	0.13	0.14	0.11	0.08
Mean		2.04	3.57	4.22	4.84	1.22	2.25	3.33	4.36	4.86	4.96

Key $\bar{x}$ mean score
 σ standard deviation
 $\sigma/\sqrt{n}$ standard error of the mean

The results for VSB noise are shown in Table 3 and Fig. 8. There is very little difference between the results for the three pictures; in particular, there is no significant difference between the moving sequence and the two still pictures so, in the future, tests on noise can be conducted using still pictures only.

The results for Test Card G show an

anomalous result at 42.2 dB. The results from the four sessions were examined separately and it was seen that two sessions gave unexpectedly low results at 42.2 dB when compared with the results for Test Card G at other SNRs. As the low results appear in two sessions, an error in setting SNR seems unlikely and the anomalous result must remain unexplained.

Table 3: Results of the second series of tests:
results for twenty non-expert observers.

Condition		White noise		De-emphasised triangular noise		VSB noise					
Picture	SNR dB	28.00	38.00	26.20	36.20	22.20	27.20	32.20	37.20	42.20	∞
Girl with Toys	$\bar{x}$	1.64	4.22	3.03	4.78	1.11	1.95	3.40	4.53	4.75	4.90
	σ	0.59	0.48	0.66	0.42	0.31	0.66	0.59	0.51	0.44	0.30
	$\sigma/\sqrt{n}$	0.10	0.08	0.10	0.07	0.05	0.11	0.09	0.08	0.07	0.05
Test Card G	$\bar{x}$	2.08	4.20	2.50	4.58	1.30	2.40	3.35	4.39	4.32	4.82
	σ	0.47	0.61	0.60	0.55	0.46	0.71	0.62	0.55	0.66	0.38
	$\sigma/\sqrt{n}$	0.07	0.10	0.09	0.09	0.07	0.11	0.10	0.09	0.10	0.06
Moving sequence	$\bar{x}$	1.87	4.26	3.12	4.68	1.08	2.05	3.59	4.37	4.71	4.90
	σ	0.52	0.55	0.61	0.47	0.27	0.71	0.61	0.56	0.46	0.30
	$\sigma/\sqrt{n}$	0.08	0.09	0.10	0.08	0.04	0.11	0.10	0.07	0.05	0.05
Mean		1.86	4.23	2.88	4.68	1.16	2.13	3.45	4.43	4.59	4.87

Key $\bar{x}$ mean score
 σ standard deviation
 $\sigma/\sqrt{n}$ standard error of the mean

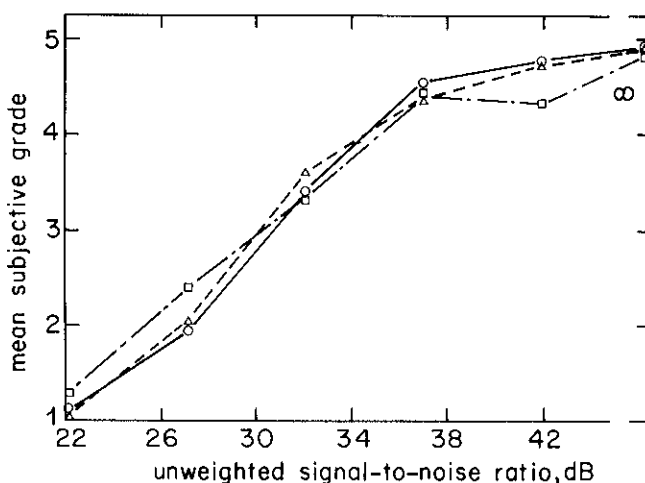


Fig. 8 - Comparison of the noise visibility for the three pictures using VSB noise.

△ --- △ Moving Sequence
 ○ --- ○ Girl with Toys
 □ --- □ Test Card G

4.3 Comparison of the results of the two series of tests

4.3.1 Consistency of results for white noise and de-emphasised triangular noise

Fig. 9 shows the mean results from the first series of tests (as given in Fig. 7), with the results of the second series of tests for white noise and de-emphasised triangular noise added. The four results from the second series of tests show strikingly good agreement with the full results of the first series of tests. This remarkable consistency means that the results of the two series of tests can be directly compared, even though the results of the first series are based on the votes of only seven non-expert observers.

4.3.2 Comparison of results for white noise, VSB noise and de-emphasised triangular noise

Fig. 10 shows the subjective test results for white noise, VSB noise, and de-emphasised triangular noise. The curve for VSB noise is very similar to the curve for white noise, but about 1½ dB more VSB

noise is needed to give the same impairment as a given level of white noise over the range 24 to 38 dB. Similarly, about 5 dB more de-emphasised triangular noise is needed to give the same impairment as a given level of white noise. These differences are analysed more accurately in Section 4.4.

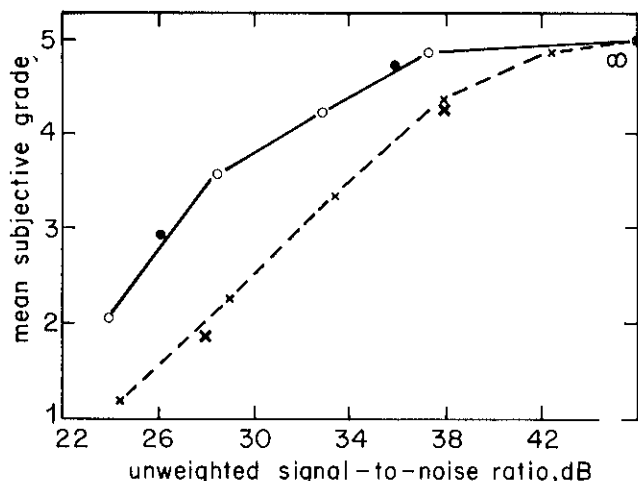


Fig. 9 - Comparison of mean results of first series of tests and second series of tests.

○ — ○ De-emphasised triangular noise } results of first series of tests
 × — × White noise }
 ● — ● De-emphasised triangular noise } results from second series of tests
 × — × White noise }

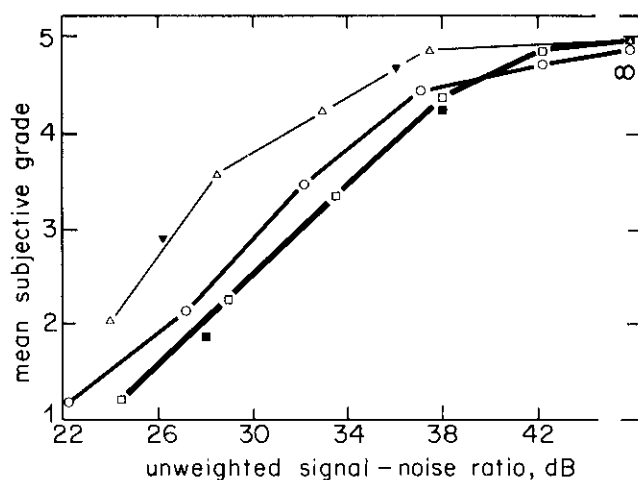


Fig. 10 - Mean subjective test results for all three forms of noise.

□ — White noise - 1986 tests
 ■ — White noise - 1987 tests
 △ — De-emphasised triangular noise - 1986 tests
 ▼ — De-emphasised triangular noise - 1987 tests
 ○ — VSB noise - 1987 tests

4.4 Further analysis of results

A method of analysis of subjective test results is suggested in CCIR Report 405-5¹⁰. A simple transformation is applied to the mean scores to give a

parameter I ; when $\log I$ is plotted against the logarithm of the objective measure of impairment (in this case, SNR) a straight line termed the impairment characteristic is obtained. Values of I are obtained as follows:

Let the mean score be U_m , which can vary between 1 and 5. The scale of values U_m is normalised by taking a continuous variable u such that

$$u = (U_m - 1) / 4$$

Then $I = (1/u) - 1$

Graphs of I (on a logarithmic scale) against SNR are shown in Fig. 11. Note that an SNR expressed in dB is already in logarithmic form, so Fig. 11 shows a linear scale of dB.

As expected, the experimental points for each form of noise lie very close to three straight lines. The straight lines were fitted 'by eye', rather than by using a mathematical criterion. It proved easy to fit three straight lines with the same slope so this was done. If the impairment characteristic were plotted with a linear vertical axis, it would appear as a sigmoid curve and the slope of the straight line in the logarithmic graph would be a measure of the steepness of the transition region of the sigmoid curve. It is not obvious that the three forms of noise should necessarily produce impairment characteristics with the same slope, but Fig. 11 shows that the slopes cannot be very different.

The presentation of results as an impairment characteristic makes it easy to derive figures for the relative effect of the different forms of noise. By measuring the horizontal distance between the impairment characteristics, it can be seen that a given level of white noise produces a subjective effect equivalent to a level of VSB noise 1.3 dB greater or a level of de-emphasised triangular noise 5.0 dB greater. This point is developed further, with reference to accepted weighting curves, in Section 5.3.

5. COMPARISON WITH OTHER RESULTS

5.1 Comparison with the results of Allnatt and Prosser

As mentioned in Section 2 above, the results of Allnatt and Prosser⁵ are those most commonly quoted in the UK. Their paper does not give the results explicitly in numerical form; it concentrates on deriving functions representing the opinion distributions they found. Figure 6 of their paper is a sigmoid curve which fits their results for white noise. Our Fig. 12 reproduces this curve superimposed on the results of our tests with white noise. Comparison shows that our

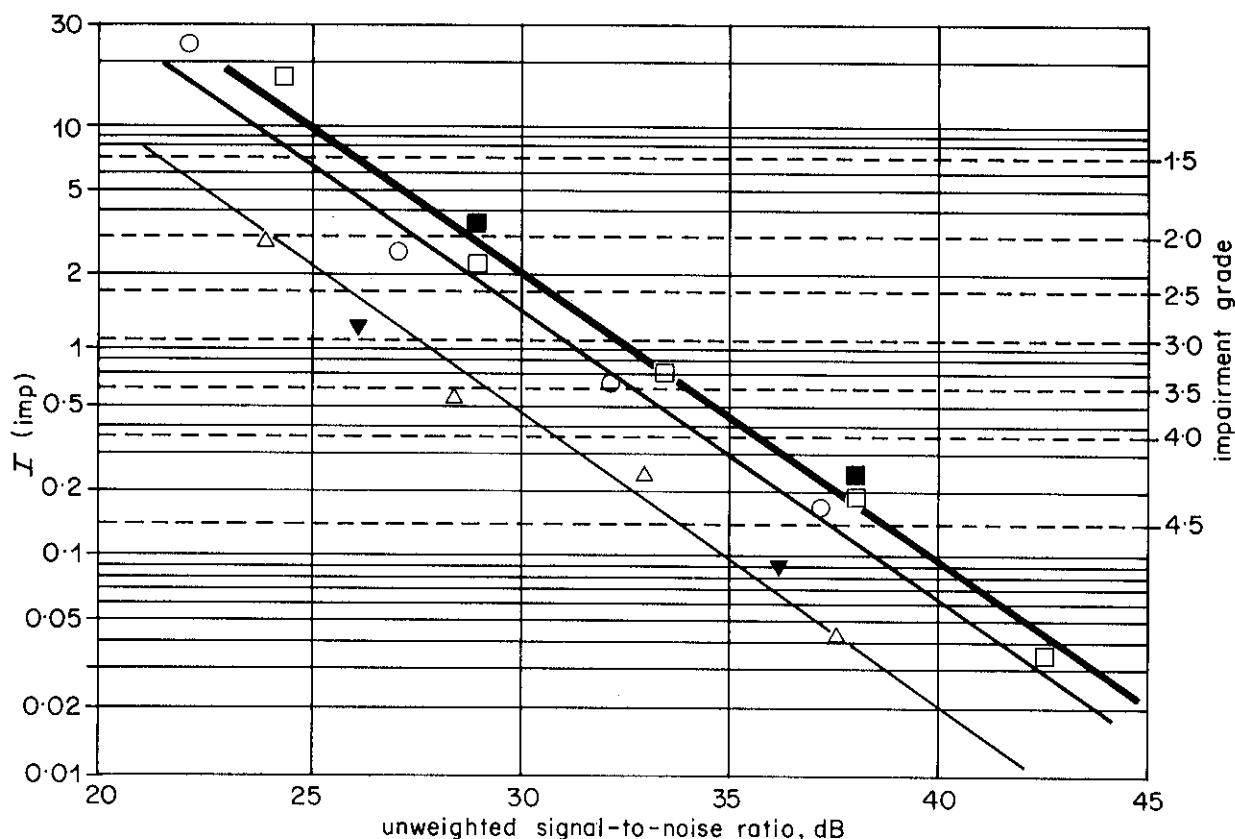
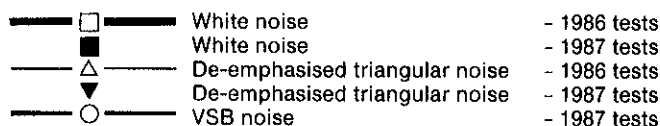


Fig. 11 - Impairment characteristics for white noise, VSB noise and de-emphasised triangular noise.



viewers found that a given level of white noise produced the same degree of impairment as a level about 2½ dB higher produced for their viewers. This difference is consistent with the suggestion in Section 2 of this Report that noise might now be more visible because of improvements in equipment and that viewers might have become more critical as they have become more familiar with good-quality colour pictures. Our results also approach the extremes of the scale (Grade 5 at high SNR and Grade 1 at low SNR) more closely than do those of Allnatt and Prosser. One possible reason for this is the use of the EBU method with a reference (unimpaired) picture in each presentation, providing better anchoring of results — Allnatt and Prosser did not use an explicit reference. Another possible reason is that the EBU method specifies the use of the CCIR 5-point impairment scale, whereas Allnatt and Prosser used the CCIR 5-point quality scale; Jones and McManus¹¹ have shown that the grade descriptions of the quality scale are less evenly spaced in the range from 'Best Imaginable' to 'Worst Imaginable' than the grade descriptions of the impairment scale.

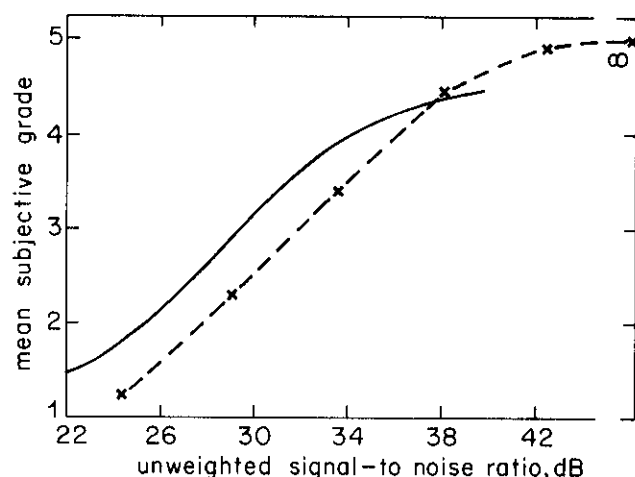
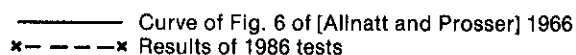


Fig. 12 - Comparison of results of 1986 tests for white noise with the results of Allnatt and Prosser (Ref. 5).



5.2 Comparison with the reference impairment characteristic of CCIR Recommendation 654

CCIR Recommendation 654¹² gives 'reference impairment characteristics' for various impairments including random noise. These characteristics 'should be regarded as a particular way of expressing the relationship between the picture quality and the objective value of each of the distortions in question, assuming that only one of them is present at any one time'. This statement does not clearly explain the purpose of 'reference impairment characteristics'. Nevertheless, it was thought desirable to compare the present results with the characteristic given in Rec. 654. Fig. 13 shows the reference impairment characteristic of Rec. 654 and also the impairment characteristic derived from the results of our tests. Our results suggest that less noise is required to produce the same impairment by about 2 dB at Grade 4.5 and about 3.5 dB at Grade 3. CCIR Recommendation 654 does not explain how the reference characteristics were derived but the characteristic for white noise may well be based partly on Allnatt and Prosser's results. The Recommendation ought, perhaps, to be reviewed in the light of our results.

5.3 Comparison with results derived from weighting curves

Noise is often measured after its spectrum has been shaped by a weighting network. The weighting networks, of which there are several in use for different purposes, usually have a greater loss at high frequencies than at low frequencies; in this way they allow for the fact that the eye is less sensitive to high-frequency noise than low frequency noise. The intention is that weighted SNRs can be compared directly, whatever the original spectrum of the noise.

For System I, two separate weighting networks were originally used, one for luminance noise and one for chrominance noise (noise in the region of the PAL subcarrier frequency). Later, Allnatt and Prosser suggested combining these two networks into a hypothetical composite weighting network (HCWN). However, in 1978 the CCIR adopted Recommendation 567 which specifies a single weighting network applicable to all systems. Inevitably these weighting characteristics are a compromise, particularly those that are agreed internationally like that of Recommendation 567. A comparison of the differences in visibility of the three types of noise tested with the

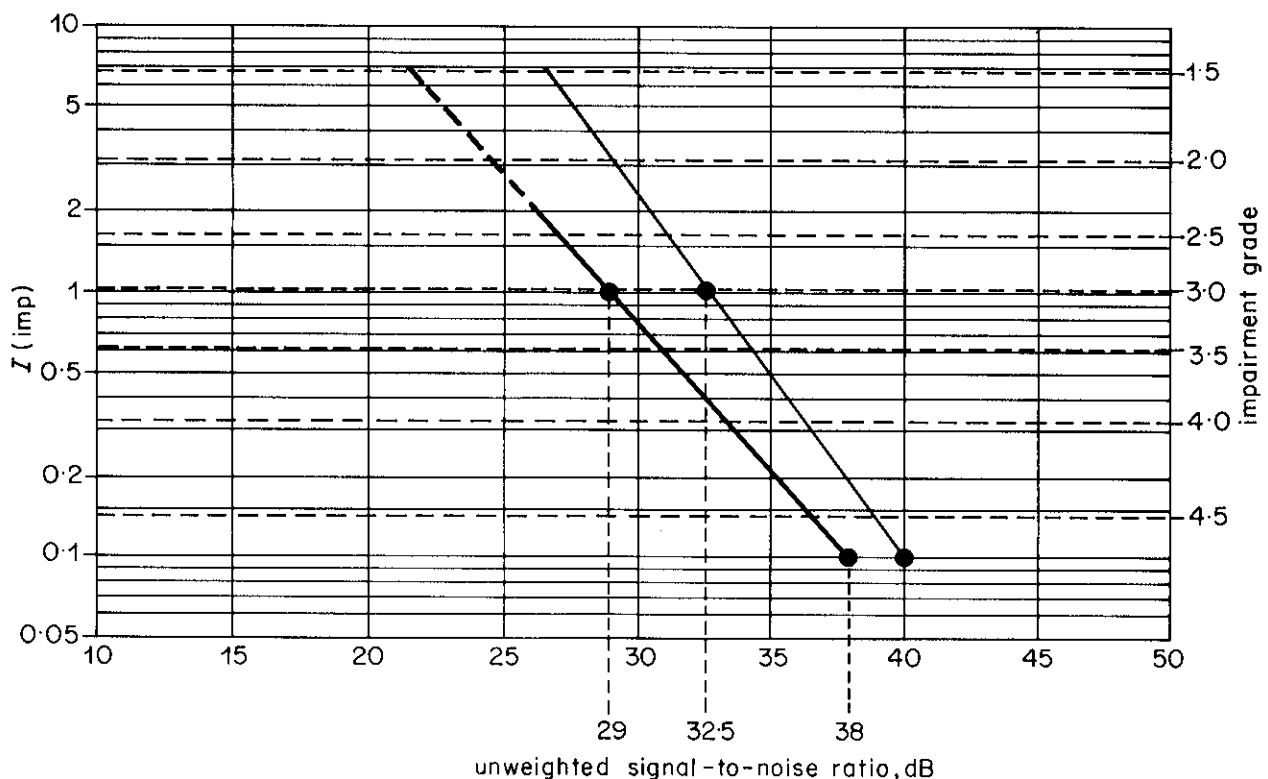


Fig. 13 - Comparison of the reference impairment characteristic given in CCIR Rec. 654 with the impairment characteristic derived from the results of the 1986 tests.

— Rec. 654
 - - - 1986 tests

differences calculated from the characteristics of a weighting network can give an indication of the validity of the weighting network.

Macdiarmid and Allnatt¹³ give a table of weighting factors for various noise spectra and weighting networks (based on calculations by Corbett¹⁴). From their table, the differences in weighting factors between white noise and the other two noise spectra tested have been calculated and are given in Table 4 along with the differences read off Fig. 11.

Table 4

	From Fig. 11	Rec. 657 weighting	System I luminance	HCWN (5.5 MHz)
VSB noise	1.3 dB	1.0 dB	1.0 dB	0.6 dB
de-emphasised triangular noise	5.0 dB	3.8 dB	4.4 dB	0.9 dB

The weighting network that agrees most closely with the experimental results is the System I luminance weighting network. The hypothetical composite weighting network clearly overestimates the visibility of chrominance noise in the conditions of this experiment; some support for this view is given by the results of the informal experiment reported in Appendix 2. However, the relative visibility of luminance and chrominance noise in PAL is a very complicated subject: it depends on factors such as the chrominance bandwidth of the decoder, the brightness and saturation of the picture, and (in a domestic receiver) the action of the automatic chrominance control. White and Reid⁶ found that the unified weighting curve of Rec. 567 underestimated the visibility of chrominance noise, whereas our results suggest the reverse. Perhaps the only conclusion that can be drawn from this discussion is that comparisons of noise visibility based on the application of weighting functions should be treated with some reserve.

6. RECOMMENDATIONS

Comparison of our results with the results of previous work suggests that observers have become more critical of noise in pictures. The tests should therefore be repeated at intervals of five years or so. The results of several such tests would be more trustworthy than the results of one isolated experiment. For a reliable comparison, at least twenty non-expert observers at a viewing distance of 6H should be used.

We have identified one anomaly in our results that deserves further investigation. Analysis of the

results of the first series of tests suggested that there was no difference in impairment seen by observers at 4H and 6H. However this result is based on the analysis of small numbers of results extracted from the total, and so cannot be considered as trustworthy. The effects of viewing distance and screen size on the visibility of noise deserve more rigorous investigation.

Such an investigation might be worthwhile in the context of HDTV. When a broadcasting format is devised for HDTV it will be necessary to determine the effect of noise on picture quality. It is difficult to carry out experiments on full-size HDTV sources and displays and results are often extrapolated from tests with smaller screens and shorter viewing distances. If such an extrapolation is carried out, then the effects of viewing distance and screen size should be investigated at the same time.

The use of weighting functions to estimate the visibility of noise with a given spectral distribution from subjective results using a different distribution should be treated with some reserve. Noise weighting functions are intended to give a rough-and-ready estimate of noise visibility for convenience in operational measurements; they should not be treated as precision instruments. This comment can be readily understood if it is remembered that noise in television, and the response of the human eye, vary horizontally, vertically, and temporally; yet the real weighting networks in use have a response only in the horizontal direction.

7. CONCLUSIONS

The tests described in this Report have given up-to-date results for the impairment caused to PAL colour television pictures by various levels of white noise, VSB noise, and de-emphasised triangular noise. The results are particularly useful in the form of the impairment characteristics given in Fig. 11. Analysis of the results has shown that a given level of white noise causes the same impairment as a level of VSB noise 1.3 dB greater or a level of de-emphasised triangular noise 5.0 dB greater.

Analysis of the results of the first series of tests found no difference between results for viewers at 4H and 6H and that expert viewers consistently rated pictures with moderate amounts of noise about a half grade lower than non-experts. The results of the second series of tests showed that movement in pictures did not mask the effects of noise.

Comparison with the results of Allnatt and Prosser⁵ suggests that observers in our tests found that a given level of white noise produced the same degree

of impairment as a level 2½ dB higher produced for their observers. Similar results are obtained from comparison of our results with the reference impairment characteristics of CCIR Rec. 654. This difference may be due to improvements in equipment (particularly display tubes) in the last twenty years, or to changes in the expectations of observers as they become more familiar with high-quality colour pictures. In any case the results presented here should not be regarded as valid for ever.

8. REFERENCES

1. WEAVER, L.E. 1959. Subjective impairment of television pictures: effect of signal-to-random noise ratio. *Electronic and Radio Engineer*, May 1959, pp 170-179.
2. BARSTOW, J.M. and CHRISTOPHER, H.N. 1962. The measurement of random video interference to monochrome and colour television pictures. *Communications and Electronics*, Nov. 1962, pp 313-320.
3. GEDDES, W.K.E. 1963. The relative impairment produced by random noise in 405-line and 625-line television pictures. *EBU Review (Technical)* No. 78, April 1963, pp 46-48.
4. PROSSER, B.A., and ALLNATT, J.W. 1965. Subjective quality of television pictures impaired by random noise. *Proc. IEE*, Vol. 112 No. 6, June 1965, pp 1099-1102.
5. ALLNATT, J.W. and PROSSER, B.A. 1966. Subjective quality of colour television pictures impaired by random noise. *Proc. IEE*, Vol. 113 No. 4, April 1966, pp 551-557.
6. WHITE, T.A., and REID, G.M. 1981. Quality of PAL colour television pictures impaired by random noise: stability of subjective assessment. *Proc. IEE*, Vol. 128 Pt. F No.4, August 1981, pp 231-235.
7. ROBINSON, A.P. 1987. The relationship between vision carrier-to-noise ratio and picture signal-to-noise ratio in a System I television receiver. BBC Research Department Report No. BBC RD 1987/23.
8. CCIR, 1986. Recommendation 500-3. Method for the subjective assessment of the quality of television pictures. *Recommendations and Reports of the CCIR*, 1986. Vol. XI-1, pp 165-173.
9. European Broadcasting Union, 1982. EBU Technical Recommendation R28. Subjective assessment method to be normally used for 625-line television pictures.
10. CCIR, 1986. Report 405-5. Subjective assessment of the quality of television pictures. Annex III. *Recommendations and Reports of the CCIR*, 1986. Vol. XI-1, pp 184-186.
11. JONES, B.L. and McMANUS, P.R. 1986. Graphic scaling of qualitative terms. *SMPTE Journal*, November 1986, pp 1166-1171.
12. CCIR, 1986. Recommendation 654. Subjective quality of television pictures in relation to the main impairments of the analogue composite television signal. *Recommendations and Reports of the CCIR*, 1986, Vol. XI-1, pp 223-224.
13. MACDIARMID, I.F. and ALLNATT, J.W. 1978. Performance requirements for the transmission of the PAL coded signal. *Proc. IEE*, Vol. 125, No. 6, June 1978, pp 571-580.
14. CORBETT, J.M. 1977. Calculation of wideband random-noise powers in television systems. *Electronics Letters*, Vol. 13, No. 18, 1st September 1977, pp 530-532.

APPENDIX 1

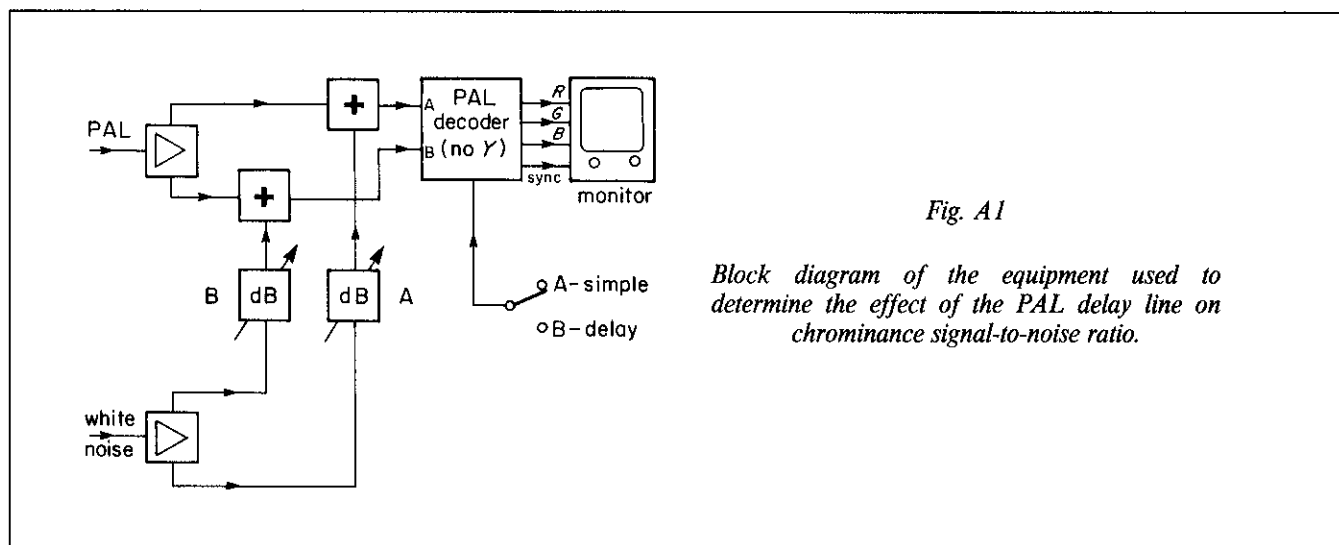
The Effect of the PAL Delay Line on Chrominance Signal-to-noise Ratio.

A1.1 Introduction

It has been assumed that a delay-line PAL decoder gives an improvement in chrominance signal-to-noise ratio of 3 dB compared with a simple PAL decoder. This is because the chrominance information from two lines is averaged, giving a 3 dB reduction in random noise. This assumption has, however, been called into question: it has been argued that although there is a reduction of 3 dB in noise power, the appearance of the noise is changed and the subjective improvement is less than the theoretical 3 dB. This appendix describes an experiment to investigate the improvement in chrominance SNR from a delay line decoder.

A1.2 The experiment

A PAL decoder was modified so that the A/B input switch and the simple/delay switch could be remotely controlled by a single switch. The luminance signal was disconnected from the output YUV-RGB matrix so that any noise on the output had come only from the chrominance circuits of the decoder. The monitor then had its brightness increased to give some luminance. Noise was added into the PAL signals as shown in Fig. A1.



It would have been better to have added noise directly to the chrominance circuits of the decoder, thus giving a clean luminance feed to the output matrix. This would have required greater modifications to the decoder; the simpler method outlined here was thought to be adequate for comparison purposes though, if absolute measurements were needed, a more elaborate arrangement would be justified.

The comparison was carried out by giving individual viewers the A/B remote control switch and attenuator B and asking them to set attenuator B so that the noise on signals A and B appeared subjectively the same. This was done for three different settings of attenuator A (29 dB, 34 dB and 39 dB). Seven engineers accustomed to judging picture quality took part in the test. The input PAL signal was 100% colour bars.

A1.3 Results

Several observers commented that the noise on the two signals was subjectively very different; this made it difficult to balance the noise levels. The table below shows the mean difference in attenuator settings (A-B) for the three noise levels, and the standard error of the mean, $\sigma/\sqrt{n}$

Attenuator A, dB	29	34	39
Mean (A-B) dB	1.54	1.14	1.07
$\sigma/\sqrt{n}$	0.15	0.26	0.20

A check with a noise meter on the output of the decoder showed the expected objective difference of 3 dB at all noise levels; the figures given above relate only to the subjective effect of the noise.

A possible reason why the difference is smaller for lower noise levels is that slight Hanover bars were visible with the decoder switched to simple PAL; the observers' eyes might follow the movement of the bars and thus tend to integrate the noise.

A1.4 Conclusions

A delay-line decoder gives a theoretical improvement in chrominance signal-to-noise ratio of 3 dB; this is confirmed by objective measurement with a noise meter. But the vertical filtering carried out by averaging across a line delay gives the noise a subjectively different character and the subjective improvement is equivalent to only 1.0 to 1.5 dB.

APPENDIX 2

An Experiment on the Relative Visibility of Luminance and Chrominance Noise*

In 1982 there was a debate on the standards to be used for DBS (Direct Broadcasting by Satellite). Frequency modulation with pre- and de-emphasis is used in the DBS channel; the noise spectrum is therefore de-emphasised triangular. It was argued by some that this would give serious impairment from chrominance noise if PAL were used for DBS. To test this argument an experiment was carried out to assess the relative visibility of luminance and chrominance noise in a PAL signal with added de-emphasised triangular noise.

A PAL decoder with A and B inputs was modified so that the luminance signal to its output matrix was derived from one input and the U and V signals from the other input. A PAL signal was fed to each of the two inputs along with variable amounts of noise as shown in Fig. A2.

In an informal subjective test, groups of engineers were asked to compare the effect of chrominance noise only and luminance noise only. The noise levels were adjusted until they judged the impairment from luminance and chrominance noise to be about equal. The absolute levels of noise were not carefully controlled but were arranged to give an impairment ranging between Grade 3 and Grade 4 on the 5-point scale. Several test patterns and EBU test slides were used as source material.

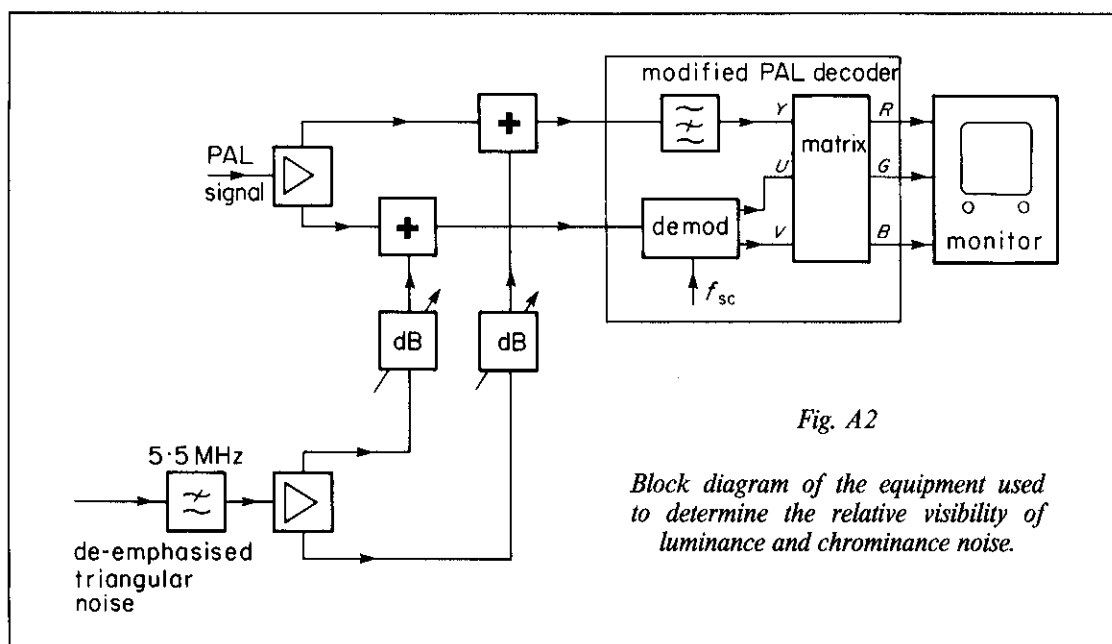


Fig. A2

Block diagram of the equipment used to determine the relative visibility of luminance and chrominance noise.

* This work was carried out by P. Shelswell.

The results are given in Table A2. In all cases but one (100% colour bars) the observers agreed that a higher level of chrominance noise was needed to give the same impairment as a given level of luminance noise.

Table A.2

Source	Amount by which chrominance noise could be increased to give equal impairment with luminance noise. (dB)
Test Card F	3
'Boats and Lighthouse'	5
'Formal Pond'	3
'Clown'	1
100% Colour Bars	-1
75% Colour Bars	0

The weighting factors given by Macdiarmid and Allnatt¹³ for de-emphasised triangular noise are 10.9 dB (System I luminance weighting) and 2.2 dB (System I chrominance weighting in 5.5 MHz bandwidth). For the same measured SNR of X dB at both inputs of the decoder, the luminance-weighted SNR is (X + 10.9) dB and the chrominance-weighted SNR is (X + 2.2) dB; that is, the chrominance-weighted SNR is lower by 8.7 dB. Now Allnatt and Prosser found equal impairment with chrominance-weighted SNR lower by only 6 dB (and this result is reflected in the hypothetical composite weighting network). We would therefore expect in this experiment that we would have had to reduce the chrominance noise (by an average of 2.7 dB) to obtain equal impairment whereas in fact it was found that chrominance noise could be increased by an average of 1.8 dB. This suggests that with modern PAL equipment chrominance noise is less important than had been believed.